

# 多人数同時発声環境における効果的なシャドーイング音声収録に関する検討\*

峯松信明 (東大), ☆楽俊偉 (東大), △山内豊 (東京国際大), △伊藤佳世子 (京大), 齋藤大輔 (東大)

## 1 はじめに

筆者らは、シャドーイング音声に対して、発話・話者単位での自動評価を検討してきた [1, 2]。先行研究では、特定の教育現場におけるシャドーイング音声を数十名の話者から収録し、自動評価技術の構築に利用した。今回、シャドーイング音声の収録を複数の研究機関から、大規模に収録する機会を得た。そこでより効果的な収録を実現すべく、収録方法（マイクの選定、収録前・収録中のインストラクションの与え方、収録ソフトの構築など）について再検討し、収録に臨んだ。本稿では、その様子について報告する。

## 2 シャドーイング音声とその特徴

シャドーイングは、モデル音声の聴取と同時にそれを追唱（即ちシャドー）する認知的負荷の高い訓練法である。元来同時通訳者の訓練法として導入されたが、現在では、外国語学習において、対象言語に対する全般的な処理能力向上を狙った訓練法として利用されている [3, 4]。対応するテキストを全く提示せずにシャドーする場合もあれば、テキストを表示したり、理解度チェックをするなど内容を理解した上で、(最後に) テキストを見ずにシャドーするなど、状況に応じて多様な利用方法が採られている。

テキスト読み上げと異なりシャドーイングは、1) 提示されたモデル音声の話速に則した発声が求められる、2) 聞き取りながら発声することが求められる、3) 個々の単語を同定できない状態でシャドーすることも頻繁にある、など、単なるテキスト読み上げに比較すれば、その認知的負荷は非常に高い。一般にシャドーイングはヘッドホンやイヤホンで音声提示され、それに応じてシャドーする訳だが「提示される音声をしっかり聞きたい」という動機から再生音量は高めになることが多く、その一方、シャドーした音声は小声となることも多い。その他シャドーイング音声特有の特徴については、[2] を参照して戴きたい。

## 3 音声収録における留意事項

### 3.1 先行研究における音声収録

筆者らの先行研究では [1, 2]、特定の教育環境（現場）でシャドーイング音声の収録を行った。規模としては 40 名程度である。Skype などのチャット用マイクとして市販されている一般的なヘッドセットマイクロホンを利用して収録が行なわれている。収録用ソフトとしては、Windows で動作するスタンドアロン型のシャドーイング音声収録ソフトを作成した。指定

したシャドーイングコンテンツ（複数のモデル音声から構成される）に対して、モデル音声を一つずつ再生し、また、同時に収録を行う。学習者のマウスクリックでモデル音声の開始と同時に収録が開始され、収録の終了も学習者がクリックで指示する、簡素なインタフェースを備えた収録ソフトである。

### 3.2 シャドーイング音声収録における問題点

先行研究における収録は、教室環境で行なわれ、語学教師と音声技術者同伴で行なわれた。収録中のトラブルは随時、学習者より教師・技術者に報告され、その都度解決に努めた。しかし、学習者自身がトラブルと認識していないことも多く、不適切な音声そのまま収録音声として提出されることも少なくなかった。今回、複数の大学・環境で大規模に収録する必要があり、収録は学習者の自宅も想定しており、語学や技術の専門家同伴での収録は不可能である。これは、一般の学習者に、収録の作法を教えつつ収録させる必要性があることを意味する。また語学教師も一般的には、収録に関する知識に乏しいため、教師・学習者両者に「収録上手」となってもらうための、教育的な音声収録インフラを構築する必要がある。

シャドーイング収録は前述したように、しっかりと聞きたい（再生音声の音量は上げたい）が、シャドー音声は小声となることが（特に初学者には）見受けられる。一般のチャット用ヘッドセットマイクを使うと、モデル音声は漏れ、それがマイクから収録されることがしばしば起こる。学習者が全くシャドーできない場合（無音）、モデル音声はその学習者の音声となって収録されるケースもあった。

更に教室環境での収録となると、多人数収録となるため、他者（自身）の声は自身（他者）にとつての雑音となる。収録音声に対して事後的に音声強調・雑音抑制の技術を適用すれば、習熟度推定精度は向上する。しかし収録時の工夫で、信号雑音比を向上できれば、それらの工夫は積極的に導入すべきであろう。

学習者の英語スピーキング能力を上げるための特効薬は「センター試験にスピーキング問題を導入することである」としばしば言われる。リスニング問題が 2006 年に導入されて 10 年経過するが、スピーキング問題は 1) 採点に時間を要する、2) そもそも回答を学生同士で発声するため試験にならない（他者に聞こえる）、などの問題がある。前者は技術的な解決を試みるのが可能だが、後者はどうであろうか？本稿では、この問題についても考察する。

次頁で対応策を示す。図・写真を効果的に利用した教示 web の URL も示すので、参照して戴きたい。

\* Experimental attempts to realize effective recording of utterances produced by multi-speakers' simultaneous shadowing. By N. Minematsu, J. Yue, Y. Yamauchi, K. Ito, and D. Saito.

## 4 各種の対応策と収録の実際

### 4.1 二つの収録環境

今回の収録は主としてK大学（以降K）の学生と、T大学（以降T）の学生を対象とする。Kは教室が利用できないため、収録は自宅で行うこととなり、Tは教室環境での収録となった。後者の場合環境雑音への配慮が必要となる。いずれの環境でも、PCノイズを拾い難いUSBマイクを使用するが、収録方向が限定される単一指向性のマイクを使用した。

### 4.2 音漏れ対策

通常のヘッドセットマイクでは、モデル音声が入る可能性がある。そこで今回の収録では、PCにUSBオーディオインタフェースを接続し、それに（カナル型）イヤホンと単一指向性マイクを接続する方法をとった。カナル型イヤホンは外耳道にイヤピースを差し込むタイプのイヤホンであり、大きな音量でモデル音声を聴取したとしても、音漏れを十分に防ぐことができる。

### 4.3 環境雑音対策

Kは個室収録、Tは教室収録であり、後者の方が環境雑音対策が必要である。他者の声を抑え、自身の声を収録する手段として（他者の口をマイクから相対的に遠ざけ、自身の口をマイクに近づける手段として）、イヤーフック型マイクを利用することとした。こうすることで、唇すぐ横にマイクを固定でき、また、ポップノイズを防ぐことも可能となる。なおKの収録は、Tとの比較、及び、費用的理由からデスクトップ型のマイクを使用することとした。この場合、ポップノイズが入らないよう、事前の注意が必要である。

### 4.4 収録前に提示する教示内容

Kでの収録であるが、近い将来、全学生に必修としてシャドーイングを課す可能性もあり、収録の不備によって不当に低い成績がつくことは避けなければならない。これは、各学習者に「収録上手」になってもらう必要があることを意味するが、その対策として、十分な収録用教示をwebで提供することとした。記述の粒度として参考にしたのは、センター試験におけるリスニング機材の教示である。マイク・イヤホンをPCに接続しても、PC側で適切に選択しないとそれらを適切に利用することは出来ない。USBマイクを使っているつもりが、実はPC内蔵マイクで収録され、電源ノイズが乗ることは少なくない。また、息が直接かかる位置にマイクを置くとポップノイズが乗る。力んだ発声をすると、波形振幅がAD変換時の最大振幅を超えてしまうこともある。これらの現象を図示しながら説明している。

- 収録環境Kの学生に対する収録前の教示  
<http://goo.gl/tXiKwA>
- 収録環境Tの学生に対する収録前の教示  
<http://goo.gl/uXr81r>

### 4.5 収録中に提示する教示内容

事前にいくら教示しても、それに従わない学習者は必ず存在する。彼らに従わせる方法として、下記の新たな教示を収録中に提示した。モデル音声の冒頭1秒は必ず無音とし、シャドーイング開始の1秒は、必ず環境雑音のみが収録されるようにした。この部分の収録音を学習者間で比較することで、環境雑音という観点で、どの学生が不適切な収録をしているのかを全学生に公開した。初回のシャドーイングの後、全学生の環境雑音レベルを計算し、そのヒストグラムを作成し、公開した。ヒストグラムの中には学生IDも明示している。他の学生と比較されることで、教示に従わない学生数を減らす試みである。

また、電源ノイズなど実際に観測された不適切な収録例についても図示しながら説明した。その上で、教示通りの収録をするよう、再度依頼した。

- 収録環境Kの学生に対する収録中の教示  
<http://goo.gl/cJ2wrV>

### 4.6 効率的な収録 web インタフェース

今回の収録はK、T両方の環境で数十名～百数十名の収録を予定しており、webを通じた収録インタフェースを構築した。シャドーイングコンテンツを選択すると、個々のモデル音声が入力され、同時に、収録も始まる。モデル音声終了後、3秒発つと自動的に収録が終了し、サーバーに音声が入力され、次の音声のシャドーイングへと移る。

- シャドーイング音声収録の簡単なデモページ  
<https://goo.gl/orjVqG>

### 4.7 スピーキング問題の導入は可能か？

上述した種々の工夫を施した上で、現在、音声収録している。イヤーフック型マイクを利用する場合、透明衛生マスクを付けることで、更に他人の音声の混入を防ぎ、自身の音声をもらさずに収録できるだろう。センター試験環境下でのスピーキング問題であるが、カナル型イヤホンを装着し、モデル音声を試験会場において完全同期して提示できれば、他者の音声を聞いてカンニングすることは不可能だろう。

## 5 おわりに

大規模なシャドーイング音声収録に対して、検討した項目について説明した。収録に協力して戴ける英語教員の方は、是非連絡して戴きたい。

謝辞 本研究は科研費 JSPS 16H03084, JSPS 16H03447, JSPS JP26240022 の支援を受けた。

### 参考文献

- [1] Luo et al., *Proc. INTERSPEECH*, 2807–2810 (2008)
- [2] 桑他, 音講論(秋), 2-5-9 (2016)
- [3] Y. Hamada, Shadowing: Who benefits and how? Uncovering a booming EFL teaching technique for listening comprehension. *Language Teaching Research* (2015)
- [4] T. Nakanishi et al., Extensive reading and the effect of shadowing. *Reading in a Foreign Language*, 23(1), 1 (2011)